

Genetic Methods used in Wildlife Populations

Peter Dratch, NPS & Tim King, USGS

"Traditional" Technologies

- Quantitative (Morphological/Meristic Variation)
- Husbandry
- Chromosomes (karyotypes)
- Protein (Allozymes, isozymes)



Genetic characterization of quantitative traits

- Single-locus characters:
 - Few are likely to be characterized in organisms of conservation interest
- Quantitative characters:
 - Laborious to separate genetic and environmental components

Pitfalls: Type I Error

- Concluding that two samples are genetically different when they are the same
- *Risk*: Conserving something redundant that we don't need

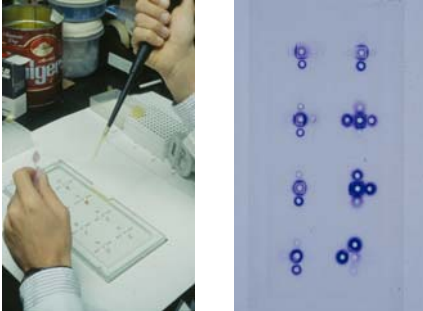
Pitfalls: Type II Error

- Concluding that two samples are genetically identical, when they are different
- *Risks*:
 - Do not conserve an important genetic component; or
 - Do not detect a change in amount of genetic variation; or
 - Send someone innocent to jail

Protein assays

- Immunology
 - useful for analysis of genetic divergence above the species level
- Protein electrophoresis
 - also called allozyme or isozyme analysis

Passive Immuno-diffusion



Allozyme variation

- Crude protein extract placed on gel material
- Protein variants are separated in electric field based on charge (also sometimes by size and shape)
- Visualized on gels with protein-specific staining system

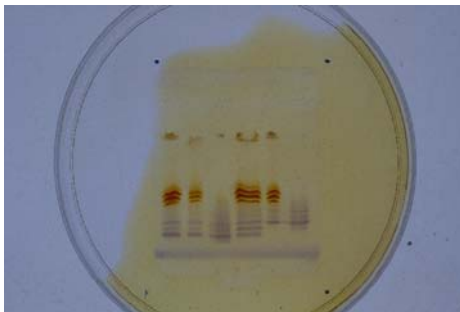
Electrophoresis in progress



Advantages of allozymes

- Inexpensive
- Models for analysis are well-developed
- Good for quick survey of general patterns of geographic variation
- Patterns revealed can be further explored using higher resolution DNA techniques.
- Process large numbers of samples quickly
- Compare data to other studies

Gpi Variation in North American Bears

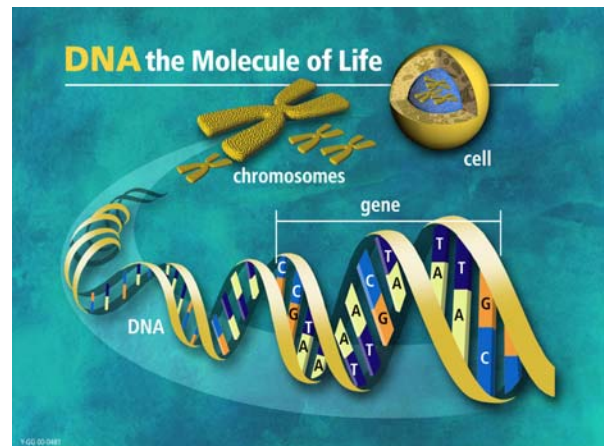


Disadvantages of allozymes

- Some groups tend to be relatively monomorphic (birds, canids, marine mammals).
- Type II errors
 - can be reduced by further analysis
 - Heat denaturation--little used
 - Two-dimensional electrophoresis
 - Sequential electrophoresis under different pH, gel concentration

Technique Chronology

<u>Time Period</u>	<u>Primary Tools</u>
1950-1970	Laboratory breeding schemes, blood types, chromosomes
1970s	Proteins (allozymes)
1980s	Mitochondrial (mt)DNA
1990s	Nuclear DNA
2000s	GenOMICS



"DNA" Technologies

PCR - polymerase chain reaction

Automated sequencing and fragment analysis

Computing

Robust statistical analyses

Internet

Genetic warehouses (GenBank, EMBL)

Bioinformatics

DNA Chips (Microarrays) and Quantitative PCR

Transgenics and other GEO methodologies

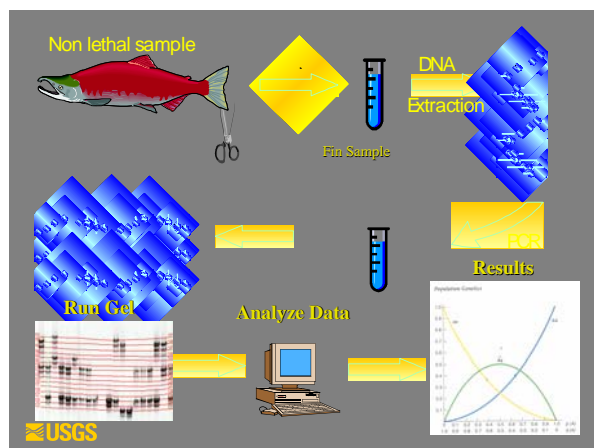
DNA-Based Methodologies

• Advantages

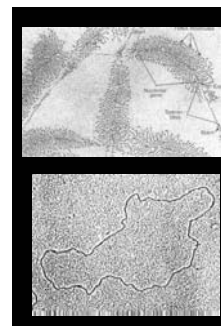
- Relative ease of collection and preservation
- Highly degraded samples are of use
- Non-invasive sampling

• Disadvantages

- Technically more demanding
- Costlier (personnel, equipment, supplies, health)



TYPES OF DNA



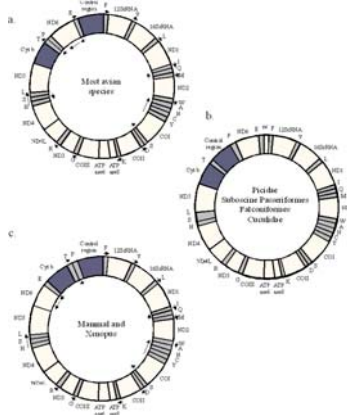
NUCLEAR

ORGANELLE

- Mitochondrial DNA (mtDNA)
- Chloroplast DNA (cpDNA)

ORGANELLE DNA

- Mitochondrial DNA
 - 15-17 KB
 - Circular
 - Eucaryotes
 - gene order varies
 - inheritance varies (usually maternal)



ORGANELLE DNA

- Chloroplast
 - 120-220kb
 - Circular
 - highly conserved
 - inheritance varies



Organelle-based Marker Systems

- Advantages
 - Rapidly-evolving relative to single-copy nuclear protein coding genes (mtDNA)
 - Easily assayed through RFLP and sequencing
 - Usually maternally-inherited: assessing bias in gene flow (mtDNA)
 - Amenable for phylogenetic and phylogeography studies

Organelle DNA-based Analyses

- Disadvantages
 - Single locus
 - Maternally-inherited
 - Rearrangements (plant mtDNA) and psuedo-genes can create problems

Genetic characterization of quantitative traits

- Single-locus characters:
 - Few are likely to be characterized in organisms of conservation interest
- Quantitative characters:
 - Laborious to separate genetic and environmental components

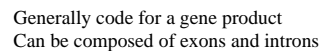
Pitfalls: Type I Error

- Concluding that two samples are genetically different when they are the same
- Risk: Conserving something redundant that we don't need

- Concluding that two samples are genetically identical, when they are different
- *Risks:*
 - Do not conserve an important genetic component; or
 - Do not detect a change in amount of genetic variation; or
 - Send someone innocent to jail

-

- Single copy regions
 - exons
 - introns
- Repetitive DNA
 - ribosomal RNA
 - SINES and LINES (Short and Long Interspersed Elements)
 - VNTRs (Variable Number of Tandem Repeats)
 - minisatellites
 - microsatellites



- tandem repeats
 - minisatellites
 - microsatellites
- interspersed repeats
 - SINES
 - LINES

Coding regions
< 5% of genome

GENE CLUSTER

GENES

MICROSATELLITE MARKERS

~95% involved in replicating DNA, gene regulation, or has no known function ("junk DNA")

USGS

Sample Preservation

Tissue Sources

- Blood
- Muscle
- Organ (e.g., liver)
- Hair follicles
- Epidermis
- Fin clips
- Scales
- Scat
- "Snail trails" and other fluids
- Bones/teeth
- Fossilized tissue

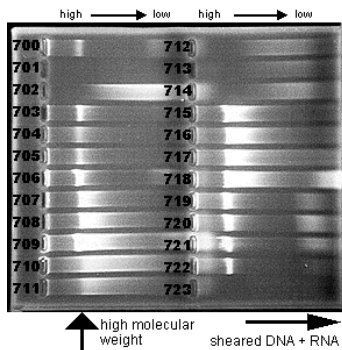
Preservation Methods

- Fresh (protein/DNA)
- Frozen (protein/DNA)
- Ethanol (DNA)
- Silica Gel (DNA)
- Filter paper (DNA)
- FTA cards (DNA)
- Buffers (DNA)
 - Longmire et al. (1988)
 - Seutin et al. (1991)
 - Urea-SDS

DNA EXTRACTION

- 1) Add a proteinase that denatures other proteins but leaves nucleic acids intact
- 2) Separate proteins from nucleic acids using organic solvents (phenol/chloroform) or high salt
- 3) Purification of DNA by alcohol precipitation or immerse tissue in solution containing detergent that lyses the nuclear membrane

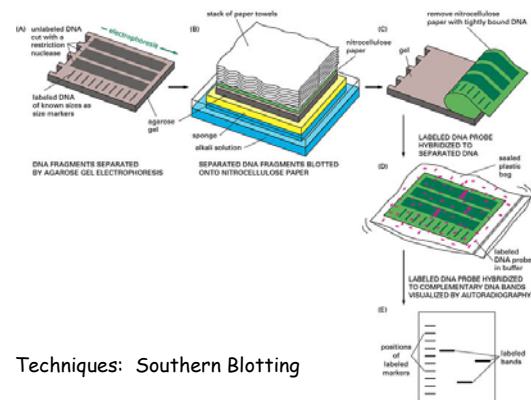
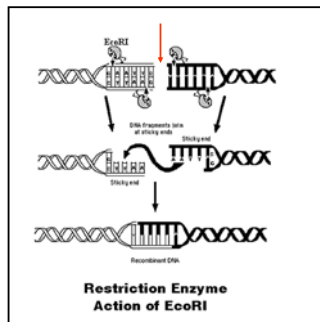
DNA QUALITY - visual inspection



VISUALIZATION Recording and Interpreting

- DNA-based
 - Dyes (Ethidium bromide/SYBR green) and transillumination, followed by photography
 - Radiation or colorimetric and x-ray photography
 - Silver Staining
 - Fluorescence and lasers

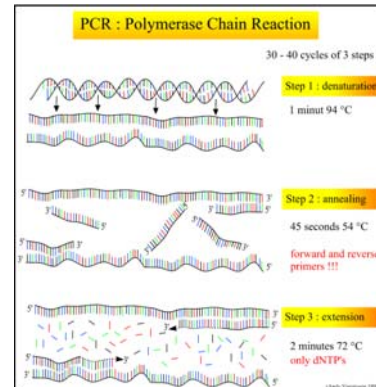
DNA Tools: Restriction enzymes



Techniques: Southern Blotting

DNA ANALYSES

PCR Sequencing SSR
 AFLPs ISSRs RAPDs Karyotyping/FISH
 QTLs Genotyping
 DGGE SNPs
 automation microarrays RFLP



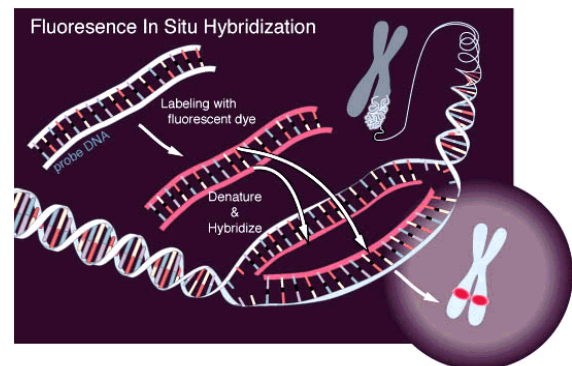
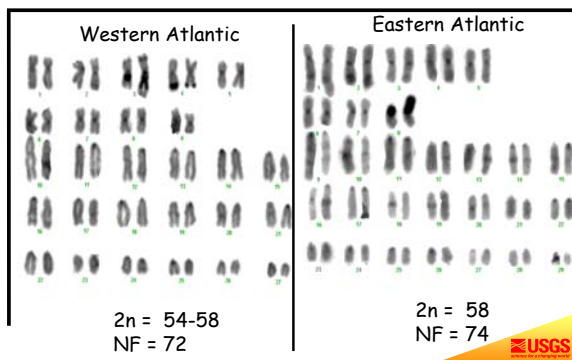
MOLECULAR MARKERS

- Variable Number of Tandem Repeats (VNTRs; "DNA fingerprinting")
 - microsatellites
 - Single-copy minisatellites
 - Multilocus minisatellites

MOLECULAR MARKERS

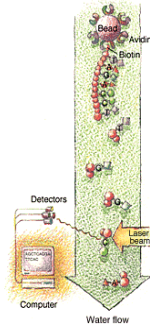
- Anonymous nuclear markers
 - Single copy nuclear genes (may include SNPs)
 - Randomly Amplified Polymorphic DNAs (RAPDs)
 - Amplification Fragment Length Polymorphisms (AFLPs)
 - Inter-simple Sequence Repeats (ISSRs)
 - Interspersed Repetitive Elements
 - SINES/LINES

Cytogenetics (karyology)



Automated Sequencing Using Fluorescence

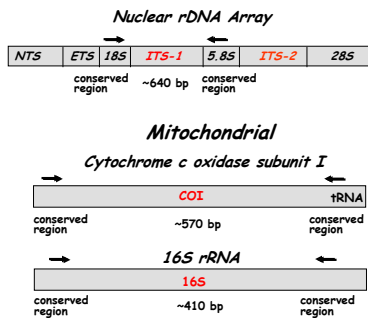
- Non-radioactive labels
 - Health and Safety benefits
- Bi-directional Sequencing
 - Quality Control
 - Time Savings
- PCR and cycle-sequencing using tailed primers
 - Cost savings
- Automated data scoring
 - Time Savings
 - Quality Control



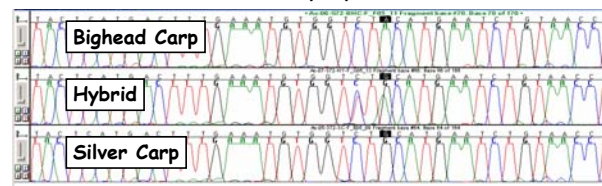
Nucleotide Sequence Information

- Advantages
 - Greatest resolution of all techniques
 - Amenable for analyses at many levels, depending upon the gene or region sequenced
 - Can tie variation to different domains or sites (more targeted analyses)
 - Variation widely distributed throughout the genome

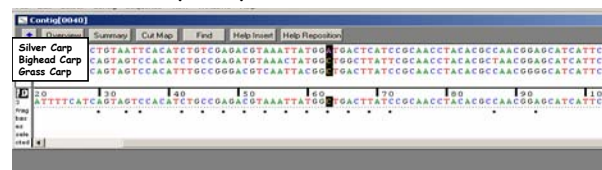
DNA Sequence Analysis



Asian Carp Hybrid Identification



Hybrid qualification - maternal contribution

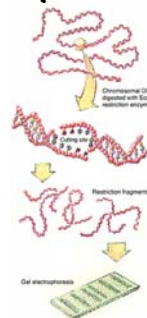


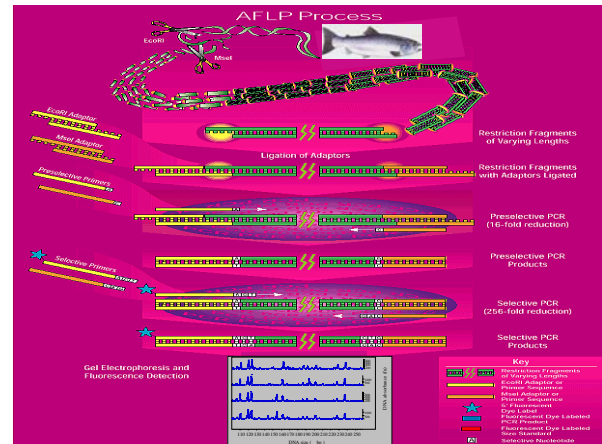
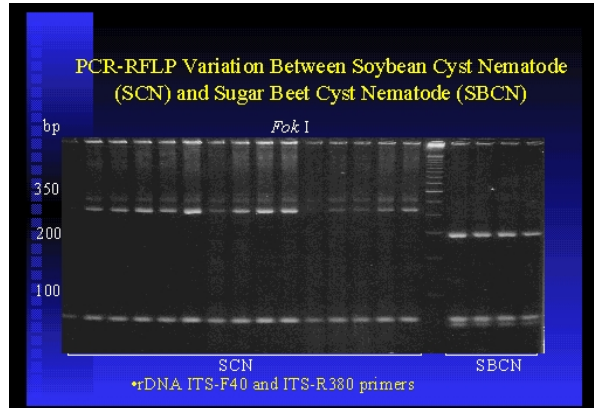
Nucleotide Sequence Information

- Disadvantages (more like hurdles)
 - Cost of reagents and identifying informative regions
 - Equipment requirements
 - Can be technically difficult

Restriction and Amplified Fragment Length Polymorphisms

- Variation
 - Presence/absence
 - Insertion/deletion (fragment size change)
- Can also be used for mutation screening





Randomly Amplified Polymorphic DNA

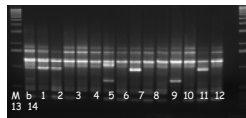


FIGURE 2. Electrophoresis gel of RAPD bands of 15 individual samples of one cultivar of canola. M, marker; b, bulk of 40 individual plants. From: M. Vondra, R.J. Maller, N. Wrotten, 2002

Variable Number of Tandem Repeats (VNTRs)

- Multilocus minisatellites
- Single locus minisatellites
 - (also seen in organelle DNA)
- Microsatellites
 - SSRs (Simple Sequence Repeats)
 - (also seen in chloroplast DNA)

VNTRs

- Advantages
 - Highly polymorphic
 - Amenable to individual-specific analyses
 - Unique multilocus genotypes
 - Amenable to population-level analyses
 - Loci are numerous and relatively evenly distributed in the genome
 - Codominant (microsats, single-locus minisatellites)
- Disadvantages
 - Many loci too polymorphic to be useful for population-level or systematic relationships
 - Can be expensive to develop and characterize

MULTILOCUS MINISATELLITES

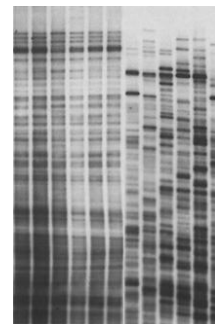


Fig. 1 DNA fingerprint patterns of brown trout produced using the Jeffreys' 33.6 human-derived probe. Fragment sizes range from approximately 2 kb (bottom) to 23 kb. Samples are (from left): three from Riabhaich, three from Cnocach, two individuals each from a fish farm and two wild populations in Ireland.

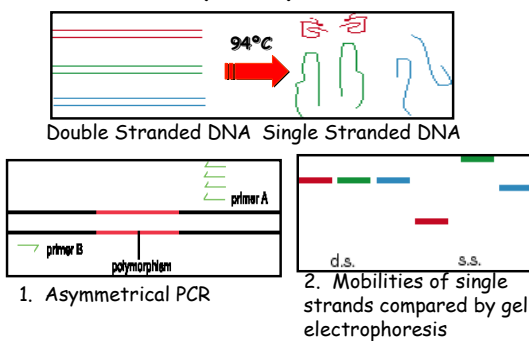
MUTATION SCREENING

- SSCP (Single-Stranded Conformation Polymorphism)
- DGGE (Denaturing Gradient Gel Electrophoresis)
- BESS T
- BIOCHIPS

Single Stranded Conformation Polymorphisms

- Codominant
- Precise knowledge of sequence polymorphism not necessary
- Strands can be distinguished by attaching different fluorescent labels to the two primers
- SNP techniques can be used if the nucleotide responsible for the mobility difference is known
- Denaturing Gradient Gel Electrophoresis (DGGE) techniques can be used for heteroduplex analysis

Single-Stranded Conformation Polymorphism



MICROSATELLITE DNA LOCI

Simple sequences, tandemly repeated

Examples: GA, CA, TA, ATT, AAT, ATC, GATC, GATA, GACA

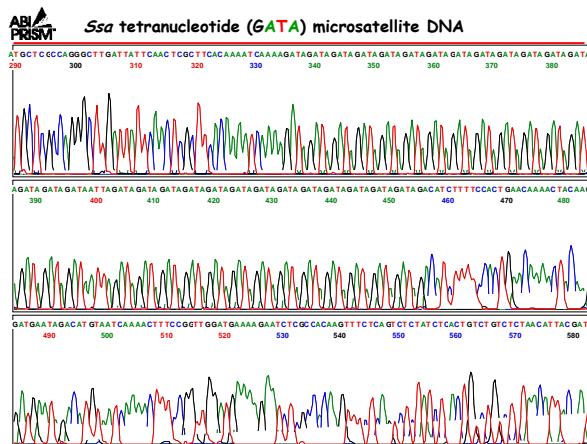
Variation measured in base pairs (i.e., numbers of repeats)

Flanking region Flanking region

GTCTAGCTAG **GATAGATAGATAGATA** TTGTCCACCTGACCTGCAATGCAT

GTCTAGCTAG **GATAGATAGATAGATAGATA** TTGTCCACCTGACCTGCAAG

GTCTAGCTAG **GATAGATAGATAGATAGATAGATAGATA** TTGTCCACCTG



Anonymous Markers RAPDs, AFLPs, ISSR's

- Advantages
 - Relatively easy to assay
 - Typically variable
- Disadvantages
 - Can't tie polymorphisms to single genes
 - Can't even be sure of genome assayed
 - Difficult to assign allelic states (not codominant)
 - Some not repeatable between laboratories or even between machines in the same laboratory

ISSRs (Anonymous Marker)

- Generally higher levels of polymorphism than RFLPs and RAPDs
 - Due to detection rather than actual polymorphism
- Error rate approximately $\frac{1}{2}$ that of RAPD markers (Yang et al. 1996)
- No prior sequence knowledge is required
- Technically simple
- Reliable once optimized

Have proven useful in plants, inverts., and birds

QTL Analysis

- The premise of QTL analysis: If a simple (qualitative) phenotype is determined by the alleles at a locus, then complex (quantitative) phenotypes must also be determined by the allelic structure at a particular locus (or loci).
- The loci determining quantitative phenotypes are thus quantitative trait loci (QTL).
- The challenge is to *find how many loci are involved, where these loci are located in the genome, and how much each allele contributes to the phenotype.*

ISSRs

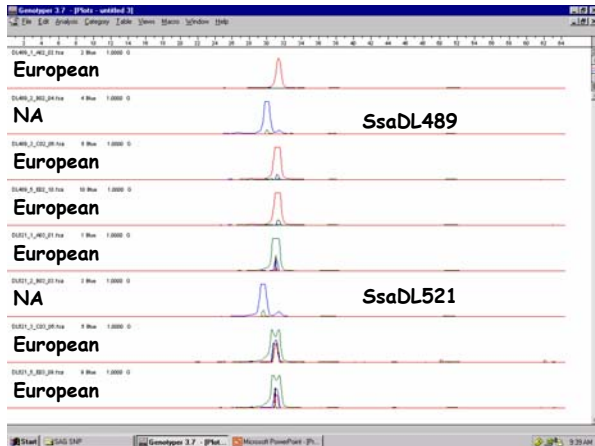
- Inter-Simple Sequence Repeats: Multilocus marker system based on the presence of VNTRs in the genome
- Primers based on a repeat sequence $(CA)_n$ are made with a degenerate 3' -anchor
- '5-CACACACACACACACACACACACACARG-3'
- Primers are used in a PCR reaction such that the reaction amplifies the sequence between two SSRs

Quantitative Trait Loci

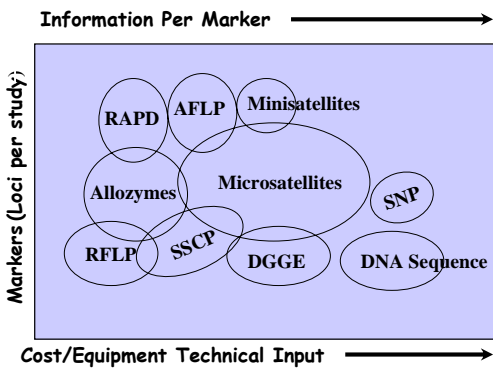
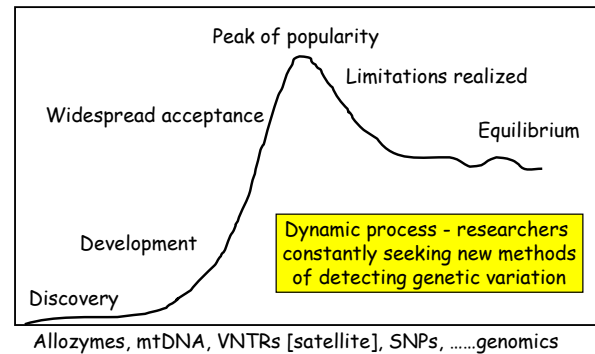
- **quantitative genetics** The area of genetics concerned with the inheritance of continuously-varying traits. Most practical domestic animal and crop improvement programs involve the application of quantitative genetics.
- **quantitative inheritance** Inheritance of measurable traits (height, weight, color intensity, etc.) that depend on the cumulative action of many genes.
- **quantitative trait** A measurable trait that shows continuous variation; a trait that can not be classified into a few discrete classes.
- **quantitative trait locus (QTL)** A locus that affects a quantitative trait. The plural form (quantitative trait loci) is also abbreviated as QTL.

Single Nucleotide Polymorphisms (SNPs)

- Genetic variation at the highest possible resolution (single base)
- Usually two alleles per locus, enabling full automation with high density DNA arrays ("gene chips")
- Numerous and widespread
- Moderately to highly polymorphic but...
- Unknown location and function, markers could be linked to each other
- Complicated technology, becoming routine



"MOLECULAR EUPHORIA"



Gabriele Procaccini, October 10, 2002; ISBW5, Ensenada, Baja California

Marker Choice

- Choose the marker
 - that exhibits the level of variation consistent with the question being asked
 - Consistent with limitations in terms of collecting and preservation of material
 - Consistent with budget in terms of cost of equipment, technical complexity, health

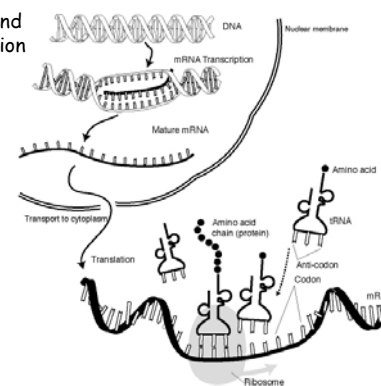
New Technologies

Microarray Technology

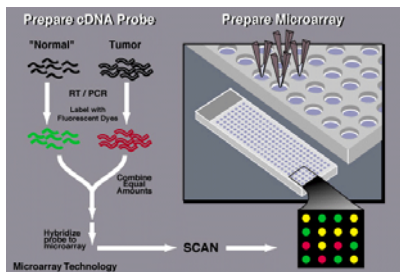
DNA chip Technology (SNPs)

Quantitative Trait Loci (QTL)

Microarrays and Gene Expression

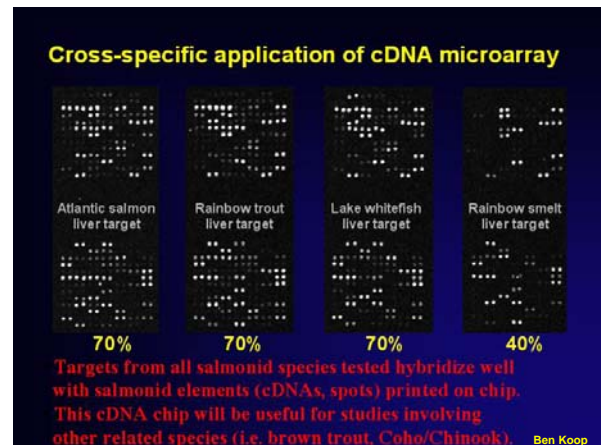
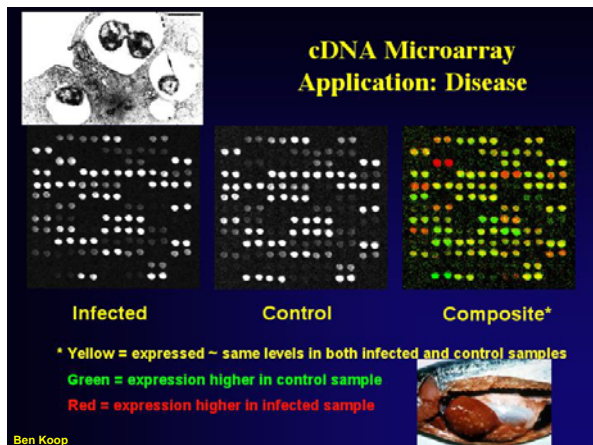


Microarray Technology



Applications: Microarrays

- Disease Diagnosis
- "Pharmacogenomics" (Drug discovery)
- Gene Discovery
- Toxicological Research "Ecotoxicogenomics"
 - Hybridization of functional genomics and molecular toxicology
 - Find correlations between toxic responses to toxicants and changes in genetic profiles of organisms exposed to toxicants



STSs and ESTs

PCR and sequencing together have made possible the creation of useful landmarks in the genome. These are several thousand short fragments of known DNA sequence whose presence in any DNA sample can be tested by PCR. They are known as **STSs** (Sequence Tagged Sites).

If an STS is part of a transcribed sequence it is known as an **EST** (Expressed Sequence Tag). Hundreds of thousands of ESTs have been created and can be accessed by computer.

Meet the 'omics'

Genome - the complete set of genes within a cell

Genomics - the genetic make-up of organisms

Transcriptome - the complete set of RNA messages coded from the DNA within the cell

Proteome - the complete set of proteins within a cell

Proteomics - the study of protein structure and function; how they interact with each other

Metabolome - the complete set of substrates and by-products of enzymatic reactions that directly impact the phenotype of the cell.

Acknowledgements

The DNA portion of this module was based on the original presentation made by Dr. Sandy Talbot (USGS, Alaska Science Center) during the seminal offering of Applied Conservation Genetics in 2002. Many of the slides used here were modifications of Dr. Talbot's original presentation.

